UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 10/757,978 | 01/14/2004 | Matthew S. Jacobsen | 23901-08207 | 4270 |

758 7590 08/18/2006

FENWICK & WEST LLP
SILICON VALLEY CENTER
801 CALIFORNIA STREET
MOUNTAIN VIEW, CA 94041

| EXAMINER | |
|---|---|
| BROWN JR, NATHAN H | |

| ART UNIT | PAPER NUMBER |
|---|---|
| 2121 | |

DATE MAILED: 08/18/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

| Office Action Summary | Application No. | Applicant(s) | |
|---|---|---|---|
| | 10/757,978 | JACOBSEN ET AL. | |
| | Examiner | Art Unit | |
| | Nathan H. Brown, Jr. | 2121 | |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *(3)* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☒ Responsive to communication(s) filed on *14 January 2004*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

4)☒ Claim(s) *1-52* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-52* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

**Application Papers**

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *14 January 2004* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *11/03/2004*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

# Examiner's Detailed Office Action

1. This Office is responsive to application 10/757,978, filed January 14, 2004.

2. Claims 1-52 have been examined.

## Claim Rejections - 35 USC § 101

3. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Claims 1-52 rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter: mathematical abstract and/or algorithm and software.

Claims 1-17 and 18 are process claims reciting a "method for selecting transformation rules for application to unstructured content" by "applying candidate transformation rules to a set of source tokens to selectively produce tokens in response to the transformation rules; determining for each candidate transformation rule a measure of accuracy of the predictive model based on actual outcomes associated with the produced tokens; and selecting transformation rules that are likely to improve the measure of accuracy of the predictive model". Clearly, these claims make only an application of mathematical abstraction and algorithm, associated with rule base

processing and predictive modeling. The claimed results do not recite any transformation of real world states. Therefore, it must be determined whether the claims produce a useful, concrete, and tangible result.

Consider the final result of the processing: "selecting transformation rules that are likely to improve the measure of accuracy of the predictive model". Here, we must question whether this result is specific, substantial, and credible. Clearly, predictive models can be used in many scientific and economic disciplines and methods of improving the accuracy of such models have as wide a range of application. However, no recitation of how the claimed invention is useful in a specific discipline is given. Further, the claim recites that the transformation rules are "likely to improve the measure of accuracy". If the utility of this result is to improve the model, then we do not know from the claim whether an improvement will occur, since the rules are only "likely" to improve the accuracy of the model. Thus, one skilled in the art of predictive modeling cannot know whether this result is substantial, as the transformation rules may or may not improve the accuracy of the model. Further, one skilled in the art cannot lend credibility to this claim without undue experimentation. Clearly, the result of claims 1 and 18 are not useful as it stands. Since the claim must be useful, concrete, *and* tangible to be statutory; the claim is not statutory.

Claims 19-35 recite a computer implemented software system, which implements the process of claims 1 and 18. Clearly claim 19 is directed to software per se. Claims 19-35 are therefore non-statutory.

Claims 36-52 recite a "computer program product, for selecting transformation rules for application to unstructured content, and storing program instructions on a computer readable medium, the instructions causing a processor to perform" the method of claims 1 and 18. We note that claim 36 has three parts: (a) the selection of transformation rules of application to unstructured content, (b) storing program instructions on a computer readable medium, and (c) instructions causing a processor to perform the method of claims 1 and 18. Now, we have shown (above) that (a) and (c) are directed toward mathematical abstraction and/or algorithm. So the question becomes: does (b) produces a useful, concrete, and tangible result.

Now, storing program instructions on a computer readable medium is clearly useful and practical for code distribution and reuse. Also, storing program instructions on a computer readable medium is routine and repeatable, thus concrete. For storing program instructions on a computer readable medium to be tangible, it must set forth a practical (or, at least, a non-abstract) application. However, we find that claim 36 does not recite a practical application of the storage step and we are also left unsure as to how the 'whole' of claim 36 functions, practically. How, practically, does (b) storing program instructions on a computer readable medium relate to (a) selecting transformation rules and (c) the instructions causing the processor to compute method of claims 1 and 18. It would seem that simply selecting transform rules for application to unstructured content would not have a causal or functional relation to storing program instructions on a computer readable medium or causing the processor to compute the method of claims 1 and 18. Without the disclosure of functional description, a computer

program product claim provides nothing more than what a program listing does, which is insufficient to show that the invention is statutory. Therefore, claim 36 does not produce a useful, concrete, *and* tangible result, so claims 36-52 are non-statutory.

## Claim Rejections - 35 USC § 112

4. The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

5. Claim 1~~, 18, 19, and 36~~ 52 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the enablement requirement. The claims contains subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention. Particularly, each claim recites "selecting transformation rules that are likely to improve the measure of accuracy of the predictive model" as a final result. But, how can it be established that the transformation rules that *will* improve the measure of accuracy? We have shown above that the breath of the invention is large, as predictive modeling is applicable in many scientific and economic problem domains. We have also show above that the nature of the claims is abstract. Now, the state of the art of statistical text mining in 2004 was still largely experimental with results shown in many different problem domains, but successful technique not largely given public exposure. The level of skill of one of ordinary skill in the art can be inferred from the degree of technicality

AA 8/16/04

of current journal papers to be that of an advanced graduate researcher. The level of predictability of predictive modeling success in statistical text mining, as shown in current technical publications, is still quite low and dependent on many factors in specific problem domains. Little direction for application of the claimed results is supplied by the Applicant and no working model is recited. The quantity of experimentation needed to set up a working application of this invention, in a particular problem domain, is not addressed in the content of the disclosure. Therefore claims 1, 19, and 36 are rejected under 35 U.S.C. 112, first paragraph.

## Claim Rejections - 35 USC § 103

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

7. Claims 1 is rejected under 35 U.S.C. 103(a) as being unpatentable over *Azari et al.*, "Actions, Answers, and Uncertainty: A Decision-Making Perspective on Web-Based Question Answering", 2003 in view of *Brill et al.*, "Data-Intensive Question Answering", 2001.

Regarding claim 1. *Azari et al.* teach a method for selecting transformation rules for application to unstructured content (*see* §3.2 Query Rewriting, "AskMSR reformulates each user question into likely substrings of declarative answers to the question. For each question, several rewrites are generated using eight rewrite heuristics.", *Examiner interprets "rewrites" as transformation rules for application to unstructured content (i.e., substrings).*), comprising: providing a set of source tokens from unstructured content (*see* above, *Examiner interprets "query words" to be source tokens.*), applying candidate transformation rules to a set of source tokens to selectively produce tokens in response to the transformation rules (*see* §3.2 Query Rewriting, "For each question, several rewrites are generated using eight rewrite heuristics.", *Examiner interprets <LEFT> and <RIGHT> to be placement tokens generated from applying candidate transformation rules (rewrite heuristics) to the source tokens.*); determining for each candidate transformation rule a measure of accuracy of the predictive model based on actual outcomes associated with the produced tokens (*see* §5.1 Understanding the Value of Queries, "In the pursuit of limiting the number of queries issued by AskMSR, we sought to replace the expert-derived heuristic functions used in AskMSR with Bayesian models that could generate probabilities of success for various rewrites. In an initial phase of analysis, we explored models that could provide a ranking of individual query rewrites."); and selecting transformation rules that are likely to improve the measure of accuracy of the predictive model (*see* §5.1 Understanding the Value of Queries, "We employed Bayesian learning procedures to generate models from a set of training cases that could be used to infer the probabilistic lift in accuracy that queries of different types would confer. Such models promised to provide a normalized

metric for ordering sets of queries by their value, providing a decision surface for deliberating about the costs and benefits of using different numbers and types of query rewrites...").

*Azari et al.* do not teach each source token associated with at least one structured content record, and each structured content record including an actual outcome. *Brill et al.* do teach each source token associated with at least one structured content record, and each structured content record including an actual outcome (*see* §3.2 N-Gram Harvesting, "Once we have obtained the set of rewrites, we submit each reformulated query to the search engine. ... The returned summaries contain the query terms, usually with a few words of surrounding context. The summary text is then processed to retrieve only strings to the left or right of the query string...", *Examiner interprets the "returned summaries" to be at least one structured content record including an actual outcome, the query answer.*). It would have been obvious at the time the invention was made to persons having ordinary skill in the art to combine *Azari et al.* with *Brill et al.* to add analysis and control of the heuristic processes to AskMSR.

8. Claim 18 is rejected under 35 U.S.C. 103(a) as being unpatentable over *Brill et al.*, "Data-Intensive Question Answering", 2001 in view of *Azari et al.*, "Actions, Answers, and Uncertainty: A Decision-Making Perspective on Web-Based Question Answering", 2003.

Regarding claim 18. *Brill et al. teach* a method for selecting transformation rules for application to unstructured content (*see* §3.1 Query Reformulation, *Examiner asserts that "only simple string matching" is used to select transformation rules.*), comprising: providing an **index of**

**source tokens** from unstructured content (*see* §3.1 Query Reformulation, "Given a query such as "Who is $w_1$ $w_2$ ... $w_n$", where each of the $w_i$ is a word, we generate a rewrite for each possible position the verb could be moved to...", *Examiner asserts that i = 1,2,...,n are indices of source tokens.*), each source token associated with structured content records (*see* §3.2 N-Gram Harvesting, "Once we have obtained the set of rewrites, we submit each reformulated query to the search engine. ... The returned summaries contain the query terms, usually with a few words of surrounding context. The summary text is then processed to retrieve only strings to the left or right of the query string...", *Examiner interprets the "returned summaries" to structured content records associated with the source tokens for each rewrite.*), each structured content record including a predicted outcome from a predictive model (see §3.2 N-Gram Harvesting, "", Examiner interprets the "returned summaries" to be at least one structured content record.); applying candidate transformation rules to the source tokens to selectively produce tokens in response to the transformation rules (*see* §3.1 Query Reformulation, "For each question, several rewrites are generated using eight rewrite heuristics.", *Examiner interprets <LEFT>, <RIGHT>, and <NULL> to be placement tokens generated from applying candidate transformation rules (rewrite heuristics) to the source tokens.*), associating each token produced by a transformation rule with structured content records associated with a source token (*see* §3.1 Query Reformulation, "["+is relative humidity", LEFT, 5]", *Examiner notes that the source tokens and the generated token, LEFT, is associated with a structured (i.e., 3-tuple) content record.*).

*Brill et al.* do not teach determining for each transformation rule a measure of the accuracy of the predicted outcomes from the structured content records associated with the tokens produced by

the transformation rule; and selecting transformation rules that improve the accuracy of predicted outcomes. *Azari et al.* do teach determining for each transformation rule a measure of the accuracy of the predicted outcomes from the structured content records associated with the tokens produced by the transformation rule (*see* §5.1 Understanding the Value of Queries, "In the pursuit of limiting the number of queries issued by AskMSR, we sought to replace the expert-derived heuristic functions used in AskMSR with Bayesian models that could generate probabilities of success for various rewrites. In an initial phase of analysis, we explored models that could provide a ranking of individual query rewrites."); and selecting transformation rules that improve the accuracy of predicted outcomes (see §5.1 Understanding the Value of Queries, "We employed Bayesian learning procedures to generate models from a set of training cases that could be used to infer the probabilistic lift in accuracy that queries of different types would confer. Such models promised to provide a normalized metric for ordering sets of queries by their value, providing a decision surface for deliberating about the costs and benefits of using different numbers and types of query rewrites..."). It would have been obvious at the time the invention was made to persons having ordinary skill in the art to combine *Azari et al.* with *Brill et al.* to add analysis and control of the heuristic processes to AskMSR.

9. Claim 19 is rejected under 35 U.S.C. 103(a) as being unpatentable over *Azari et al.*, "Actions, Answers, and Uncertainty: A Decision-Making Perspective on Web-Based Question Answering", 2003 in view of *Oracle*, "Oracle9i Data Warehousing Guide Release 2 (9.2): §20 OLAP and Data Mining", 1996, 2002 and further, in view of *Dittrich et al.*, "The Active Database Management System Manifesto: A Rulebase of ADBMS Features", 1995.

Regarding claim 19. *Azari et al.* teach a computer implemented software system for selection of content transformation rules (*see* §3.2 Query Rewriting, "AskMSR reformulates each user question into likely substrings of declarative answers to the question. For each question, several rewrites are generated using eight rewrite heuristics.", *Examiner interprets "rewrites" as transformation rules for application to unstructured content (i.e., substrings).*), the system comprising: an index of source tokens derived from unstructured content (*see* §3.1 Query Reformulation, "Given a query such as "Who is $w_1$ $w_2$ ... $w_n$", where each of the $w_i$ is a word, we generate a rewrite for each possible position the verb could be moved to...", *Examiner asserts that i = 1,2,...,n are indices of source tokens.*), each source token associated with structured content records (*see* §3.2 N-Gram Harvesting, "Once we have obtained the set of rewrites, we submit each reformulated query to the search engine. ... The returned summaries contain the query terms, usually with a few words of surrounding context. The summary text is then processed to retrieve only strings to the left or right of the query string...", *Examiner interprets the "returned summaries" to structured content records associated with the source tokens for each rewrite.*); a predictive model, adapted to generate the predicted outcome for a structured content record (*see* §5.1 Understanding the Value of Queries, "In the pursuit of limiting the number of queries issued by AskMSR, we sought to replace the expert-derived heuristic functions used in AskMSR with Bayesian models that could generate probabilities of success for various rewrites. In an initial phase of analysis, we explored models that could provide a ranking of individual query rewrites.", Examiner notes that the AskMSR predictive model generates a predictive outcome for the set of 3-tuples containing the rewrites (see

above).); and a rules selection process, adapted to apply selected transformation rules to the index to produce tokens from the source tokens, and identify transformation rules likely to improve the accuracy of the predictive model (*see* §5.1 Understanding the Value of Queries, "We employed Bayesian learning procedures to generate models from a set of training cases that could be used to infer the probabilistic lift in accuracy that queries of different types would confer. Such models promised to provide a normalized metric for ordering sets of queries by their value, providing a decision surface for deliberating about the costs and benefits of using different numbers and types of query rewrites...").

*Azari et al.* do not teach a database of structured content records, each content record including a predicted outcome; or a database of content transformation rules, each transformation rule adapted to produce a token in response to a source token. *Oracle* does teach a database of structured content records, each content record including a predicted outcome (*see* §Scoring, "ODM takes a table as input. It scores every record, and returns a scored table as a result. ... ODM can deliver a variety of scores. It can return a rating or probability of a specific outcome. Alternatively it can return a predicted outcome and the probability of that outcome occurring."). It would have been obvious at the time the invention was made to persons having ordinary skill in the art to combine *Azari et al.* with *Oracle* to support many different types of analysis in large data warehouse environments.

*Azari et al.* do not teach a database of content transformation rules, each transformation rule adapted to produce a token in response to a source token. However, *Dittrich et al.* do teach a database of content transformation rules, each transformation rule adapted to produce a token in response to a source token (*see* pp. 2-4, §2 Terminology, "ECA-rules (event-condition-action rules) consist of events, conditions and actions. ... Once a set of rules has been defined, the active database system monitors the relevant events. In this way, it detects the occurrence of each relevant event and afterwards it notifies the component responsible for rule execution about this occurrence. We call this notification the signalling of the event. Consecutively, all rules which are defined to respond to this event are triggered (or fired) and must be executed. Rule execution incorporates condition evaluation and action execution. ... An action formulates the reaction to an event and is executed when the rule is triggered and its condition holds. An action may contain data modification ... operations", *Examiner interprets both a source token and a token produced in response to a source token to be "data items that conditions and actions of rules associated to the event can refer to".*). It would have been obvious at the time the invention was made to persons having ordinary skill in the art to combine *Azari et al.* with *Dittrich et al.* to obtain a system able to react automatically to situations in a database and beyond.

10. Claim 36 is rejected under 35 U.S.C. 103(a) as being unpatentable over *Azari et al.*, "Actions, Answers, and Uncertainty: A Decision-Making Perspective on Web-Based Question Answering", 2003 in view of *Martin* (USPN 5,937,192) and further, in view of *Brill et al.*, "Data-Intensive Question Answering", 2001.

Regarding claim 36. *Azari et al.* teach a computer program product, for selecting transformation rules for application to unstructured content (*see* §3.2 Query Rewriting, "AskMSR reformulates each user question into likely substrings of declarative answers to the question. For each question, several rewrites are generated using eight rewrite heuristics.", *Examiner interprets "rewrites" as transformation rules for application to unstructured content (i.e., substrings).*), comprising: providing a set of source tokens from unstructured content (*see* above, *Examiner interprets "query words" to be source tokens.*); applying candidate transformation rules to a set of source tokens to selectively produce tokens in response to the transformation rules (*see* §3.2 Query Rewriting, "For each question, several rewrites are generated using eight rewrite heuristics.", *Examiner interprets <LEFT> and <RIGHT> to be placement tokens generated from applying candidate transformation rules (rewrite heuristics) to the source tokens.*); determining for each candidate transformation rule a measure of accuracy of the predictive model based on actual outcomes associated with the produced tokens (*see* §5.1 Understanding the Value of Queries, "In the pursuit of limiting the number of queries issued by AskMSR, we sought to replace the expert-derived heuristic functions used in AskMSR with Bayesian models that could generate probabilities of success for various rewrites. In an initial phase of analysis, we explored models that could provide a ranking of individual query rewrites."); and selecting transformation rules that are likely to improve the measure of accuracy of the predictive model (*see* §5.1 Understanding the Value of Queries, "We employed Bayesian learning procedures to generate models from a set of training cases that could be used to infer the probabilistic lift in accuracy that queries of different types would confer. Such models promised to provide a

normalized metric for ordering sets of queries by their value, providing a decision surface for deliberating about the costs and benefits of using different numbers and types of query rewrites...").

*Azari et al.* do not teach a computer program product for storing program instructions on a computer readable medium, the instructions causing a processor to perform the operations. However, *Martin* does teach a computer program product for storing program instructions on a computer readable medium, the instructions causing a processor to perform the operations (*see* Fig. 4 and col. 5, lines 55-56, *Examiner interprets "storage area 310" to be a computer readable medium.*). It would have been obvious at the time the invention was made to persons having ordinary skill in the art to combine *Azari et al.* with *Martin* to allow execution on a multiple distributed computer system.

*Azari et al.* do not teach each source token associated with at least one structured content record, and each structured content record including an actual outcome. *Brill et al.* do teach each source token associated with at least one structured content record, and each structured content record including an actual outcome (*see* §3.2 N-Gram Harvesting, "Once we have obtained the set of rewrites, we submit each reformulated query to the search engine. ... The returned summaries contain the query terms, usually with a few words of surrounding context. The summary text is then processed to retrieve only strings to the left or right of the query string...", *Examiner interprets the "returned summaries" to be at least one structured content record including an actual outcome, the query answer.*). It would have been obvious at the time the invention was

made to persons having ordinary skill in the art to combine *Azari et al.* with *Brill et al.* to add analysis and control of the heuristic processes to AskMSR.

## Correspondence Information

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Nathan H. Brown, Jr. whose telephone number is 571-272- 8632. The examiner can normally be reached on M-F 0830-1700. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Anthony Knight can be reached on 571-272-3687. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306. Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Anthony Knight
Supervisory Patent Examiner
Tech Center 2100

Nathan H. Brown, Jr.
August 16, 2006